October 4, 2010

EX PARTE NOTICE

Electronic Filing

Ms. Marlene H. Dortch
Secretary
Federal Communications Commission
445 12th Street SW
Washington, D.C. 20554

RE: **In the Matter of Implementing a Nationwide, Broadband, Interoperable Public Safety Network in the 700 MHz Band**
PS Docket No. 06-229

Dear Ms. Dortch:

On October 1, 2010, the undersigned met with the following members of the Public Safety and Homeland Security Bureau (PSHSB): Jamie Barnett, Bureau Chief; David Furth, Deputy Bureau Chief; Jennifer Manner, Deputy Bureau Chief; Jeff Cohen, Senior Legal Counsel; Genaro Fullano, Associate Chief; Jason Kim; Brian Hurley; Roberto Mussenden and Behzad Ghaffari. Also present, from the Office of Engineering and Technology (OET), was Walter Johnston, Electromagnetic Compatibility Division Chief and Michael Ha; and from the Wireless Telecommunications Bureau (WTB) was Tom Peters, Chief Engineer.

The meeting took place to discuss technical and policy aspects of public safety roaming and priority access to LTE-based wireless networks as set forth in the attached paper (also available at http://www.ece.cmu.edu/~peha/priority_roaming_for_public_safety.pdf).

Pursuant to Section 1.1206(b)(2) of the Commission's rules, an electronic copy of this letter is being filed with the office of the Secretary.

Respectfully submitted,

/s/Ryan Hallahan
Ryan Hallahan
Hallahan@cmu.edu
Ph.D. Student
Department of Engineering and Public Policy
Carnegie Mellon University

Policies for Public Safety Use of Commercial Wireless Networks[*]

Ryan Hallahan[†] and Jon M. Peha[‡]
Carnegie Mellon University

Abstract

Wireless broadband functionality could revolutionize the way public safety responds to emergencies by bringing capabilities to first responders they have never before had access to. By also providing public safety users with roaming access to commercial networks, on a priority basis, it is possible to increase the aggregate capacity, geographic coverage, and service reliability available to these users far beyond what would be possible with networks dedicated to public safety users alone. The benefits of roaming depend in part on how this priority access is structured. This paper studies the technical and operational issues associated with providing public safety users with priority roaming access. Results are applicable in any nation where public safety users have access to both public safety and commercial networks that are based on the Long Term Evolution (LTE) technology standard. First, the paper shows that the LTE standard offers a wide range of priority-related capabilities that could be important for making roaming valuable to public safety users. These include the ability to differentiate traffic streams based on a number of factors that reflect their relative importance, to manage traffic such that objectives with respect to data rate and quality of service (QoS) can be met for the important streams, and to make more resources available for more important streams during times of heavy utilization by either blocking, dropping, or lowering the QoS of streams that have been identified as less important. Second, this paper addresses some of the issues associated with potential agreements among the many public safety agencies and commercial operators. Such agreements may prove important in ensuring that the needs and expectations of all parties are met. This paper describes how guidelines that promote commonality can reduce transaction costs, but shows how such agreements should also consider and accommodate some forms of technical diversity that can emerge across implementations, even when they all comply with the LTE standard. Third, this paper discusses issues associated with making a given standard such as LTE useful for public safety purposes even as both the standard and the needs of public safety evolve over time. Finally, this paper shows that LTE supports a wide range of operational policies and arrangements, allowing policy makers to consider important non-technical factors when choosing among these arrangements. One option is an automated system, in which decisions that affect priority are made based on factors that the network knows *a priori* (e.g. user identity) or that can be detected by the network automatically (e.g. roaming status). Alternatively, the system might allow real-time human intervention so that priority-related decisions can be based in part on situational information that the system cannot detect on its own, such as the intention of users. This paper shows that the LTE standard could support a range of technical implementations, in which different degrees of responsibility for these real-time interventions can be assigned to individual first responders, to more centralized representatives of public safety, or to operators of commercial networks.

[*] The authors gratefully acknowledge the financial support of the John and Claire Bertucci Fellowship.

[†] Ryan Hallahan, Ph.D. Student in the Department of Engineering and Public Policy, Carnegie Mellon University, hallahan@cmu.edu, www.contrib.andrew.cmu.edu/~rhallaha/

[‡] Jon M. Peha, Professor of Electrical Engineering and Public Policy, Carnegie Mellon University, peha@cmu.edu, www.ece.cmu.edu/~peha

All research presented here was performed at Carnegie Mellon University. Opinions expressed belong to one or both authors, and do not represent any other individual or organization.

1. Introduction

Wireless broadband networks present a unique opportunity to revolutionize the way public safety responds to emergencies, bringing a number of new and important applications to first responders who previously had to rely on only narrowband voice (Peha, 2005; 2007a). In order to bring this functionality to first responders, public safety agencies around the world may deploy wireless broadband networks (Hallahan & Peha, 2008; 2009; 2010; Peha, 2008). In addition to public safety-deployed networks, commercial operators will continue to deploy and operate their own commercial broadband systems. Public safety access to these commercial networks can act as a valuable supplement to the services provided by a dedicated public safety infrastructure. But traditionally, public safety agencies have been unable to rely on commercial networks for mission-critical communications.[1] If, however, public safety users can roam onto commercial networks seamlessly (i.e. using the same devices they use on their own networks), and receive preferential treatment when they do (i.e. have priority access to resources when the network is congested), then public safety may be able to rely on this access more than ever before.

Given the unique and vital nature of public safety's mission, many issues, such as technical and operational design decisions, should be addressed so that public safety is better able to rely upon commercial networks to meet public safety's critical communications. The technical design decisions include determining how to utilize mechanisms in the wireless broadband technology to treat public safety users in a preferential manner, such that their access to commercial networks can be relied upon even in times of emergency. The operational design decisions include determining how to utilize mechanisms in the wireless broadband technology to change the preferential treatment public safety users receive in response to the current situation by enabling humans to intervene. In this paper, we study the issues and tradeoffs presented by specific technical and operational design decisions as well as the policy implications these decisions may have, especially on the agreements that will govern public safety priority roaming access to commercial broadband networks.

Providing public safety users roaming access to commercial wireless broadband networks can yield a number of benefits including: (1) increased aggregate capacity, (2) increased coverage, (3) increased resiliency, and (4) increased cell site diversity. More specifically, where commercial wireless service is available in addition to public safety wireless service, public safety will have access to increased aggregate capacity to support their activities when their own

[1] This is due, at least in part, to the fact that public safety and commercial systems have, typically, been designed for different needs and used incompatible technologies. Historically, in the U.S. and many other nations, public safety agencies and commercial carriers have built and operated their own separate wireless communication networks. Commercial carriers have typically built their networks to cover larger regions using commercial technology standards with performance and features targeted at consumers. Meanwhile, public safety agencies have typically designed systems to serve their own agencies and area of jurisdiction, using specialized technologies designed to meet the mission critical needs of emergency responders (usually public-safety-grade land mobile radio (LMR) systems) (Hallahan & Peha, 2008; Peha, 2006; 2007b).

networks are fully loaded.[2] In addition, where commercial wireless service is available but public safety service is not, public safety will have access to increased geographic coverage by roaming onto commercial networks in these areas. Also, where commercial wireless service is available in addition to public safety service, public safety will have access to more resilient, fault-tolerant, and dependable communications capabilities in the event that something happens to their own network infrastructure (and this resiliency increases as more common points of failure are separated). Finally, where commercial wireless service is available in addition to public safety service, but their cell sites are not co-located, public safety users may have access to additional cell sectors and, in the best case, cell sectors with which they can connect with greater signal strength compared to connecting to their own network (either due to distance or terrain between user and cell site). This may enable either a higher data rate transmission than would otherwise be possible or the same data rate but using fewer spectrum resources (i.e. greater spectral efficiency) which, effectively, would provide greater aggregate capacity.

Public safety can realize some benefits of roaming access whether or not they receive preferential treatment on commercial networks. However, by providing preferential treatment to public safety users, the degree to which public safety users can rely on commercial roaming increases, and thus the benefits derived from having roaming access increases. In fact, recently, priority roaming access for public safety has received increased attention in the US due, at least in part, to the Federal Communication Commission's (FCC) National Broadband Plan (NBP). The NBP recommends that public safety be able to roam with priority access onto commercial networks so that, when necessary, commercial networks can supplement the service provided by public safety's own networks (Federal Communications Commission [FCC], 2010a). However, many of the specific details of the proposed roaming with priority access were left to be worked out.

In this paper, we study roaming with priority access when there can be several broadband networks run by different entities (both commercial and public safety), where both commercial and public safety users could be trying to access the same cell sites, with many different applications in use (each with different bit rate and quality of service (QoS) requirements), and in any given location multiple cell sites or cell sectors may be within range of a single user. There can be emergencies in which congestion occurs and many different users with different credentials using different applications will be competing for the same resources. The priority roaming scheme should be able to coordinate activity in times of congestion (i.e. give

[2] Compared to commercial systems, it is extremely difficult to determine what the worst-case load will be on a public safety system in any given year, but, given the mission of public safety agencies, meeting this worst-case load is very important. At the same time, the peak to average ratio of load on a public safety system is much higher than on a commercial system, meaning that if public safety designs its network to meet the absolute worst-case load scenario, much of the capacity on their network will lay idle most of the time. Doing so can substantially increase the cost of the public safety network (Hallahan & Peha, 2008; 2010), but makes inefficient use of that infrastructure. On the other hand, if public safety can make use of commercial networks in addition to their own in times of extreme emergency, it is more likely they will be able to support the increased load in these instances.

preferential treatment to some requests over others) so that the right users and applications receive the right resources at any given time. To do so, we study wireless broadband networks based on the Long Term Evolution (LTE) technology standard. These assumptions would apply to the US if the recommendations of the NBP are adopted and given that relevant wireless broadband networks (i.e. networks with which priority access roaming is likely to be involved) are likely to use LTE for their deployments (e.g. public safety regions that have received waivers to begin deploying networks (Federal Communications Commission [FCC], 2010b) and several major commercial operators (3G Americas, 2010)).

In particular, we analyze the roaming and priority mechanisms available in LTE, determine the functionality they provide, identify potential technical and policy issues related to these mechanisms, and determine possible modifications or additions to the standardized mechanisms. One decision that must be made is whether to adopt a purely automated priority system (i.e. a priority system based on predefined rules that makes decisions using what the network either knows or can detect about public safety's instantaneous QoS requirements), or one where greater situational awareness is supported through some form of human intervention. We will discuss technical design decisions, initially assuming a system in which LTE's mechanisms are used to meet public safety's needs in an automated approach. Then, when discussing operational design decisions, we will study how, in a LTE-based network, human intervention in priority decisions can accommodate additional functionality that may be beneficial to public safety beyond what a purely automated priority system can provide (i.e. affecting priority decisions by having humans intervene to provide situation-specific information that the network cannot detect on its own), and some of the benefits and complications that come with such an approach.

By providing specific examples of technical and operational design decisions, this paper will also identify how these decisions affect other issues such as the agreements between public safety and commercial operators. A key to these agreements is determining what needs to be specified for consistency across commercial providers and across time periods, balancing the desire to enable innovation over time and diversity of offerings across providers while ensuring there remains sufficient commonality to provide the required functionality. For example, agreements will need to be able to handle possible technological developments which may be necessary either now or in the future if, for instance, specific standardized mechanisms in LTE need to be modified to meet the needs of public safety.

Section 2 of this paper provides some useful background and context by analyzing the existing priority access policy for the wireless voice networks in the US and by highlighting the differences between preferential treatment on a voice-only network versus on a broadband data and voice network. Section 3 provides an overview of a LTE-based network and highlights the important mechanisms and concepts in the LTE standard that enable preferential treatment of users and applications. Sections 4 and 5, respectively, analyze the technical and operational design decisions that may need to be made to ensure that a commercial network meets the needs of public safety, and identify the policy implications these decisions may have. Finally, section 6 discusses the conclusions of this paper.

2. Policies for Priority Access to Voice-only versus Broadband Systems

In this section, we will investigate the differences between priority access policies in a voice-only network versus priority access policies in a broadband data and voice network. To provide some background on the topic, section 2.1 will discuss the existing priority access policy for the wireless voice networks in the US and identify the relevant design decisions. Then, section 2.2 will identify the differences between preferential treatment on a voice-only network versus a broadband data and voice network.

2.1. The Wireless Priority Service

In the US, the existing cellular voice infrastructure has been equipped with priority access system for national security and emergency preparedness users over the last decade. This system, the Wireless Priority Service (WPS), was designed to provide only voice priority access. In the US, the WPS system represents the current state-of-the-art in public safety priority access to commercial wireless networks and, thus, serves as a useful introduction to the topic of public safety priority access to commercial broadband networks. By identifying the design decisions made in the WPS policy, we will provide context for the design decisions we discuss in sections 4 and 5.

The framework of the WPS was presented in a petition by the National Communications System (NCS) (1995), and the subsequent Report and Orders from the FCC (1998; 2000). In these reports, the FCC outlined the permissible parameters for a priority system: the wireless priority service must be invoked on a call-by-call basis, it must be available for the national security and emergency preparedness (NS/EP) user community, it must provide five levels of prioritization, it must be voluntary on the part of the carriers, and carriers must provide reasonable capacity for the public even though they are providing a priority service for the government.

To invoke a priority call in the WPS system, authorized users must dial a special standardized code before the number they are dialing. The destination of that call can be either a wireless or wire line subscriber, and it isn’t necessary for that subscriber to be a WPS authorized user. If, due to congestion, the call cannot be completed initially it is placed in a queue at both the originating and terminating points in the network. As resources become available on the network, they are awarded first to users in the priority queue in accordance with their predetermined level of priority access.

Thus, the WPS system provides NS/EP users with priority without preemption, which means that idle resources go to highest priority call. However, when all voice channels are busy, a high-priority call must wait in a queue until an existing call ends. This design means that there is a risk that high priority calls may experience significant delays, although the availability of a queue means that the call does not need to be retried repeatedly.

Users who wish to become authorized for WPS access must apply for service and, if approved, they are assigned one of five priority levels, with Priority Level 1 having the highest priority access. The five currently established NS/EP priority levels are: Priority 1, executive leadership and policy makers;[3] Priority 2, disaster response and military command and control;[4] Priority 3, public health, safety, and law enforcement command;[5] Priority 4, public services/utilities and public welfare;[6] and Priority 5, disaster recovery[7] (National Communications System). In the WPS system, all five of these priority levels are prioritized over commercial subscribers. This design means that the system can, at best, distinguish 6 classes of resource requests.

It is up to individual users to determine when it is appropriate to invoke priority access since WPS is always available for authorized users, even if no emergency is taking place. Thus, in the WPS system, there are no limits on how often an individual user may place priority calls.

However, the WPS system does place a limit on the number of voice channels that can be used for priority calls at any given time. This limit is 25% of voice channels on each cell site at any given time (Ackerman, 2003; Chambers & Riley, 2004). When WPS usage reaches the 25%-priority-utilization limit, NS/EP users form their own priority queue to access that 25% of capacity as it becomes available. This design ensures that priority usage does not take up a disproportionate amount of capacity. This means commercial users are not starved of access to the network, even if there are many priority users requesting resources from the network. However, this design means there is a risk that when an existing call ends, the available channel is allocated to a non-priority user while a high-priority call waits in a queue.

In all, there are about 100,000 NS/EP users currently enrolled in the WPS system, the majority of which are federal users (National Communications System, 2009). There are roughly a couple million first responders in the US and many more if other emergency responders and related public safety personnel are included (Hallahan & Peha, 2008; 2009; 2010). Thus, the WPS system serves a relatively small fraction of the overall public safety community in the US.

[3] Priority 1 is meant for users at the highest level of government (all levels: federal, state, and local), such as the President and Cabinet members; State Governors and Cabinet-level officials; and Mayors and County commissioners; and some members of each of their respective senior staff.

[4] Priority 2 is meant for users that are key to managing the initial response to an emergency at the local, state, regional and federal levels, such as emergency directors and incident command managers from each level of government.

[5] Priority 3 is meant for personnel who direct operations critical to life, property, and law and order, such as law enforcement, emergency medical services, and fire and rescue leaders and commanders at each level of government.

[6] Priority 4 is meant for users whose responsibilities include managing public works and utility infrastructure restoration.

[7] Priority 5 is meant for individuals who are responsible for managing various recovery operations after the initial response has been accomplished.

User agencies are responsible for paying for their standard wireless handset and subscription costs plus some additional monthly and usage-based WPS specific charges[8] (all payments are made to the wireless carriers) (National Communications System). Meanwhile, the NCS program covers all costs for the infrastructure enhancements required to make the networks capable of supporting the WPS system. Thus, the costs associated with the deployment of the priority system are borne by the federal government while the costs of using the service are passed to local, state and federal agencies through usage-based charges. This design means that those users who use priority resources more also end up paying more.

2.2. Preferential Treatment in Voice versus Broadband Systems

By providing preferential treatment to public safety users on commercial networks, it is possible to increase the degree to which these users can rely on this access. This is the motivation behind the WPS system providing public safety access to the commercial voice systems in the US, as we discussed in section 2.1. There are, however, distinct differences between providing preferential treatment to traffic in a voice-only network versus a broadband data and voice network. In this section, we will clarify the differences between the two and highlight the implications this has on a priority access policy for a broadband data and voice network.

Generally speaking, in a multi-channel, circuit-switched voice network with several users, there are usually many, fixed size channels which are either being used or not depending on how many users are making calls. In such a system, it is straightforward to discuss traditional mechanisms such as priority and preemption[9] that enable preferential treatment of certain calls over others. Essentially, the only choices to make are: if all channels are in use and additional users are requesting resources, either a given call can be preempted (resulting in a dropped call for the preempted user and the channel being allocated to the user requesting the resource), or not and the requesting user has to either wait in some queue or (if no queue is present) be blocked and have to retry their call attempt. If queues are present, then channel requests can be queued based on the priority of the request (so that when a channel becomes available it can be assigned to the highest priority user requesting it), or not and the queue is simply filled and serviced in order of arrival with no requests receiving preference over another. While it is possible to add additional complexity using these priority and preemption mechanisms, such as reserving some capacity for a subset of priority users or creating complicated rules governing who is allowed to preempt whom, the fundamental mechanisms of priority and preemption are all that is needed.

However, things can be quite different in a packet-switched broadband voice and data network which supports many different applications with diverse and variable resource usage. Resource

[8] The WPS specific charges include a onetime WPS activation fee of up to $10, a monthly WPS service feature cost of no more than $4.50, plus a usage fee of no more than a $0.75 per minute when WPS is invoked (i.e. by dialing *272).

[9] There are several ways to define priority and preemption policies. In this work we use the term preemption to refer to mechanisms to stop another user's use of spectrum resources (also referred to as *ruthless preemption*). Whereas we use the term priority to refer to a variety of mechanisms to allocate resources to one user over another when both request access to the same resource or to give one user preferential placement in a queue for spectrum resources before other users.

usage in a broadband network is, ultimately, dependent on which (of the numerous available) applications are being used. Each application may have different QoS requirements, and some may require a minimum guaranteed bit rate while others can tolerate a non-guaranteed bit rate, best-effort service. Thus, designing a priority policy is much more complicated than simply deciding which user should be prioritized over another or who should be allowed to preempt others.

Even in a broadband network, it may be possible to treat real-time services (which require constant, guaranteed bit rates) in the same way that voice calls are treated in a circuit-switched voice-only network (i.e. using simple priority and preemption mechanisms). (Although even simple prioritization rules would likely become more complicated as they would need to factor in not only which user was requesting resources, but also which application is being used for every combination of user and application.) However, real-time, constant bit rate services are only one part of the equation. For other real-time services, it may be possible to throttle back their usage and reallocate those resources to other users and/or applications.

There are also non-real-time, best effort services which have QoS requirements that are very different from real-time services (e.g. no need for a guaranteed bit rate). These applications, such as web browsing and email access, are bursty in nature and can tolerate variations in the amount of resources allocated to them. For this reason, it is difficult to apply traditional terms like preemption to these types of services because there is no connection to preempt. Instead, preferential treatment of these types of services should be governed by mechanisms such as scheduling algorithms and queue management techniques. For instance, in a wireless broadband network, spectrum resources will be allocated on short, fixed intervals, and the algorithm that schedules which services are allocated what resources can do so at the packet-level, and can consider a variety of factors including various QoS requirements of the associated service (e.g. packet loss error rate and packet delay budget) and even the relative priority of the service.

Thus, when analyzing preferential treatment in a packet-switched broadband network, considering only priority and preemption mechanisms is insufficient. Instead, a variety of additional priority mechanisms and discrimination techniques may be appropriate. In the next section, we will introduce the various mechanisms available to provide preferential treatment in a LTE-based network and section 4 will describe the associated technical design decisions.

3. Overview of Important LTE Network Concepts and Mechanisms

In order for a commercial network to adequately meet the needs of public safety users, a number of technical and operational design decisions will likely need to be made. In this particular paper, we focus on technical issues, such as how roaming and QoS mechanisms are used to support public safety's QoS requirements, and operational issues, like how individuals can intervene to affect the QoS public safety users experience when public safety users are roaming on a commercial network. As an example, we study the LTE technology standard and the mechanisms it provides for roaming and QoS control because in the US both commercial networks and public safety networks will likely be LTE-based. This section serves as an introduction to the relevant LTE concepts and mechanisms that will be discussed in greater detail in sections 4 and 5. This section will first present an overview of the LTE standard, then discuss the fundamental QoS concepts, and finally discuss the QoS policy control mechanisms available.

3.1. Overview of LTE

LTE, or Long Term Evolution, refers to the Release 8 iteration of the 3rd Generation Partnership Project's (3GPP) mobile network technology family which also includes the GSM/EDGE and UMTS/HSxPA standards releases (3GPP, 2010). There are two main components of a LTE network: the radio access network and the packet core network.   LTE's radio access network is called Evolved UMTS Terrestrial Radio Access Network (E-UTRAN) and LTE's new packet core is called the Evolved Packet Core (EPC). Both of these components were designed to ensure that LTE is a packet-switched, all-IP standard in contrast to previous voice-centric, circuit-switched architectures (Johnson, 2010). (Indeed, in LTE, voice traffic will primarily be supported using Voice over IP (VoIP) technology.)

The E-UTRAN has two main elements: the User Equipment (UE) and the E-UTRAN base station (eNodeB or eNB). The UE is a generic term for the handsets and other devices that subscribers use to communicate with the eNodeB's over the network's allocated spectrum. The eNodeB handles all radio access related functions and each eNodeB communicates with the packet core. Service providers can have their own separate core networks, but share eNodeB's, since each eNodeB can be connected to multiple cores. UEs identify their service provider upon connecting to an eNodeB and then are connected to the correct core network (Johnson, 2010).

Each packet core or EPC will typically include the following network elements: a Serving Gateway (S-GW), a PDN Gateway (PDN-GW), a Policy and Charging Rules Function (PCRF), and a Mobile Management Entity (MME) (Johnson, 2010). The MME manages mobility, UE identity, and several aspects of security. The PCRF detects service flows and enforces the charging policy. The S-GW acts as a border between the RAN and the core network, managing user plane mobility and maintaining data paths between the eNodeB and the PDN Gateway. The PDN-GW provides connectivity between the core network and the external packet data networks (PDN). (PDN is just a generic term for any number of packet-based networks the EPC may be connected to, including the internet or a private or corporate PDN.)

The following figure is a generic representation of the LTE network architecture. This figure, based on diagrams in (Olsson, Sultana, Rommer, Frid, & Mulligan, 2009), includes the main elements of an LTE network and shows their general relationship to one another.

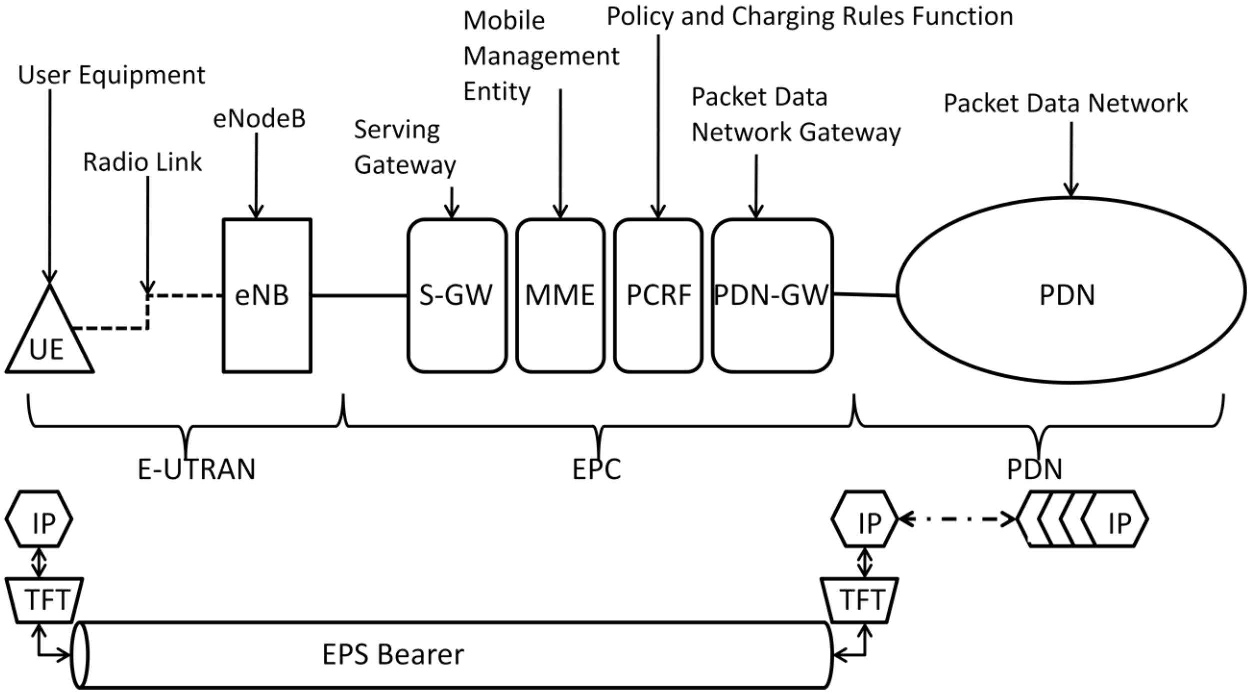


Figure 1: The main elements of an LTE network and their general relationship to each other

3.2. Fundamental QoS Concepts in LTE

The LTE standard provides a robust set of technical mechanisms which enable different users and services to receive preferential treatment (both in terms of priority access and preemption capability) across the E-UTRAN and EPC. The fundamental concepts that are central to this access and therefore we discuss in this section are: the EPS bearer (a logical channel for data flows that require the same QoS treatment), the parameters that differentiate bearers based on their QoS requirements (e.g. the Quality Class Identifier and Allocation and Retention Priority), and the numerous mechanisms which facilitate the usage of these parameters (Johnson, 2010).

Overview of EPS Bearers

At a high level, a bearer is the term used to describe a 'virtual' channel established between the endpoints of the network (i.e. from the UE to the PDN-GW). A bearer is 'virtual' in the sense that all traffic from the user device is carried across the same physical channel (the radio channel) back to the network, but many virtual channels can be created to distinguish between how different traffic should be treated over the same physical channel. There are two types of EPS bearer[10] in LTE: default and dedicated. Every UE has at least one default bearer which is established when the UE first attaches to the network and remains available for the duration of the connection. A UE can have anywhere from zero to several dedicated bearers established at any given time and each is set up and taken down on an as-needed basis.

[10] In this paper, when we only use the term 'bearer' we are, in fact, referring to an EPS bearer.

Dedicated bearers are used when the QoS requirements for some traffic is different than the QoS provisions provided by the default bearer. Furthermore, all traffic requiring the same QoS-level treatment will be carried on the same bearer (e.g. if a device is making a voice call and streaming video at the same time, and both applications require the same level of QoS, the traffic from both will be mapped to the same bearer). In this way, the bearer forms the fundamental unit for discussing the QoS mechanisms available in an LTE network. And as we will discuss below in greater detail, there are a number of ways to distinguish one bearer from another including: the guarantees and caps placed on the bit rates associated with a bearer, the preference given to a particular bearer when it is established, and the QoS guarantees given to a particular bearer at the packet level. The following table provides an overview of the types of bearer and the bit rate and QoS treatment parameters available to each (Johnson, 2010).

		Type of Bearer	
		GBR	Non-GBR
Bit Rate Parameter			
GBR:	Guaranteed Bit Rate	**X**	
MBR:	Maximum Bit Rate	**X**	
APN-AMBR:	APN Aggregate Maximum Bit Rate		**X**
UE-AMBR:	UE Aggregate Maximum Bit Rate		**X**
QoS Parameter			
QCI:	Quality Class Identifier	**X**	**X**
ARP:	Allocation and Retention Priority	**X**	**X**

Table 1: The bit rate and QoS treatment parameters available to each of type bearer

Overview of Traffic Flow Template (TFT)

Since a separate EPS bearer is established for each flow of packets that requires a different level of QoS and a single UE may have multiple bearers associated with it, a given user's IP packets (both entering the network and being transmitted by the UE) must be filtered into the appropriate EPS bearers. In LTE, this is accomplished using a Traffic Flow Template (TFT). There is an uplink TFT (UL-TFT) used by the UE, and a downlink TFT (DL-TFT) used by the PDN-GW, associated with each bearer. At a high level, the TFT is just a list of source/destination IP addresses and TCP/UDP port combinations[11] that identify which IP packets (based on their header information) should be assigned to which bearer (Olsson, Sultana, Rommer, Frid, & Mulligan, 2009). TFTs are established (or modified) at the same time that the bearer they correspond to is established (or modified).

Overview of Bit Rates for Bearers

[11] Additional parameters in the IP flow may be used in a packet filter, including (note that not all attributes may be used in the same packet filter, some are mutually exclusive): Remote IP Address and Subnet Mask; Protocol Number (IPv4)/Next Header (IPv6); Local Port Range; Remote Port Range; IPSec Security Parameter Index (SPI); Type of Service (TOS) (IPv4)/Traffic class (IPv6); Flow Label (IPv6).

For dedicated bearers, there are two possible types: guaranteed bit rate (GBR) bearers and non-guaranteed bit rate (non-GBR) bearers. For a GBR bearer, the system guarantees a minimum bit rate will be provided to that bearer once it is established. This means that GBR bearers sending at a bit rate less than or equal to their GBR can assume that packet drops as a result of congestion will not occur. For a GBR bearer, a maximum bit rate (MBR) may also be specified which caps the maximum bit rate that bearer will receive (Johnson, 2010).

For non-GBR bearers, no minimum bit rate is guaranteed by the network. Thus, there are no guarantees as to the amount of traffic a non-GBR bearer can support at any given time, which could potentially result in packet loss during times of congestion. In addition, non-GBR bearers for the same device may be capped in the aggregate bandwidth they receive by using the aggregate maximum bit rate (AMBR) parameter. The AMBR can be specified at either the APN level (APN-AMBR) or the UE level (UE-AMBR) (Johnson, 2010). For example, the UE-AMBR can be used to cap the aggregate bit rate that is allocated to all non-GBR bearers used by a given UE.

The decision of which type of bearer should be used (GBR vs. non-GBR) depends upon the service that is carried by that bearer. As discussed by Olsson et al. (2009), GBR bearers are typically used for services where it is better to block them initially rather than degrade the service after it has already started. For example, it may be desirable to block a real-time voice call before it begins during times of congestion, rather than admit the service and then have the voice be unintelligible since the guaranteed bit rate cannot be maintained. (Note that many real-time services can actually adapt to the available bit rate to some degree, but there is still a minimum bit rate below which they cannot operate properly.) On the other hand, non-GBR bearers are typically used for applications such as web-browsing and email, as these applications do not require a guaranteed bit rate. However, simply because an application does not require a GBR does not make it less important than applications that require a GBR; the relative importance of applications can depend on a number of additional factors. As discussed by Olsson et al. (2009), the choice of which bearer to use for each service is up to the operator and their configuration.

Overview of Bearer Level QoS Parameters

To support QoS requirements, the EPS bearer includes several parameters which dictate the preferential treatment a bearer may receive. Each bearer, including both GBR and non-GBR bearers, is associated with the following bearer level QoS parameters: the QoS Class Identifier (QCI) and the Allocation and Retention Priority (ARP). The QCI parameter dictates the packet-level preferential treatment a bearer receives, while the ARP parameter dictates the preferential treatment individual bearer receives when they are being established. These parameters may be specified independently of the other, allowing for many different QCI+ARP combinations for each bearer.

Overview of Allocation and Retention Priority (ARP)

When bearers are being established (or modified) on the network and resources are limited, the network may need to make decisions regarding which bearer requests should be accepted and which should be rejected (this usually occurs when available radio capacity is limited and typically involves GBR bearers). The primary role of the ARP parameter is to facilitate this decision making process (3GPP, 2008b). To do so, the ARP parameter contains three components: a single scalar value and two separate flag values. The scalar value contains information about the priority level of a bearer, while the two flags refer to the preemption capability and preemption vulnerability of the bearer.

The ARP priority level is used to ensure that the request of the bearer with the higher priority level is given preference over lower priority bearers. During periods where resources are limited, the network may choose to drop bearers of low priority to free up required resources. The pre-emption capability flag defines whether or not a given bearer is allowed to preempt (i.e. force the system to drop) other bearers of lower priority level. On the other hand, the preemption vulnerability flag defines whether or not a given bearer is susceptible to preemption (i.e. being dropped) even by bearers with a higher ARP priority level.

It should be noted that once successfully established, a bearer's ARP value has no effect on the packet forwarding treatment (e.g. scheduling and queue management) a bearer receives at a node. Indeed, the ARP is not included within the EPS QoS profile that is sent to the UE (3GPP, 2008b).

Overview of Quality Class Identifier (QCI)

Once bearers are established using the access control mechanisms provided by the ARP parameter, the nodes in the network still need to know how to treat the packets for each bearer. During times of congestion, bearers (who have been established) will compete for limited resources. This means that at individual nodes (e.g. eNodeB), the limited resources need to be allocated to individual packets from many different bearers. The QCI parameter tells the nodes how to prioritize those resources among the packets.

The QCI parameter is specified by a simple scalar value. There is one-to-one mapping of standardized QCI values to standardized QoS characteristics. The table below summarizes the QCIs that have already been standardized including: their priority level, packet delay budget, packet error loss rate, and examples of services which will typically be mapped to each QCI (3GPP, 2008a). Thus, the QCI parameter is used by the eNodeB to determine the packet forwarding treatment of each bearer (e.g. scheduling weights and queue management thresholds). This treatment is pre-configured by the operator owning the access node (e.g. eNodeB), such that the QoS requirements associated with a given QCI are met (3GPP, 2008b).

Resource Type	QCI	Priority	Packet Delay Budget[12]	Packet Error Loss Rate[13]	Example Services
GBR	1	2	100 ms	10^-2	Conversational Voice
	2	4	150 ms	10^-3	Conversational Video (Live Streaming)
	3	3	50 ms	10^-3	Real Time Gaming
	4	5	300 ms	10^-6	Non-Conversational Video (Buffered Streaming)
Non-GBR	5	1	100 ms	10^-6	IMS Signalling
	6	6	300 ms	10^-6	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
	7	7	100 ms	10^-3	Voice, Video (Live Streaming), Interactive Gaming
	8	8	300 ms	10^-6	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
	9	9	300 ms	10^-6	QCI typically used for the default bearer of a UE/PDN

Table 2: The Standardized QCI Values and their Standardized QoS Characteristics

3.3. Overview of Policy Control and Roaming

The LTE standard also provides several features for controlling and initiating the QoS mechanisms discussed in the previous section. In this section, we will first discuss the policy and charging control framework in LTE and the network elements required to support this framework, we will then compare network-initiated QoS control to terminal-initiated QoS control, and finally, we will discuss how QoS policies can be controlled when users roam on to other networks.

Overview of Policies and Charging Control (PCC)

Policy and Charging Control (PCC) is the concept in LTE that enables flow-based policy control (e.g. QoS management) and charging control (Olsson, Sultana, Rommer, Frid, & Mulligan, 2009). The main component of this concept is the Policy and Charging Rules Function (PCRF)

[12] "A delay of 20 ms for the delay between a PCEF and a radio base station should be subtracted from a given PDB to derive the packet delay budget that applies to the radio interface. This delay is the average between the case where the PCEF is located "close" to the radio base station (roughly 10 ms) and the case where the PCEF is located "far" from the radio base station, e.g. in case of roaming with home routed traffic (the one-way packet delay between Europe and the US west coast is roughly 50 ms). The average takes into account that roaming is a less typical scenario. It is expected that subtracting this average delay of 20 ms from a given PDB will lead to desired end-to-end performance in most typical cases. Also, note that the PDB defines an upper bound. Actual packet delays - in particular for GBR traffic - should typically be lower than the PDB specified for a QCI as long as the UE has sufficient radio channel quality." (3GPP, 2008a) at Table 6.1.7 Note:1.

[13] "The rate of non congestion related packet losses that may occur between a radio base station and a PCEF should be regarded to be negligible. A PELR value specified for a standardized QCI therefore applies completely to the radio interface between a UE and radio base station." (3GPP, 2008a) at Table 6.1.7 Note:2.

which is an optional element in the LTE architecture responsible for providing policy control decision and charging control functionalities that are enforced by the Policy and Charging Enforcement Function (PCEF). (Where a policy is just a set of rules that determines how a specific IP flow is treated and the QoS it receives.) The Application Function (AF) interacts with application level signaling and extracts session information that it provides to the PCRF, while the Subscription Profile Repository (SPR) contains subscription and policy information for individual users. The following figure, based on diagrams in (Olsson, Sultana, Rommer, Frid, & Mulligan, 2009), illustrates the relationship between these network elements.

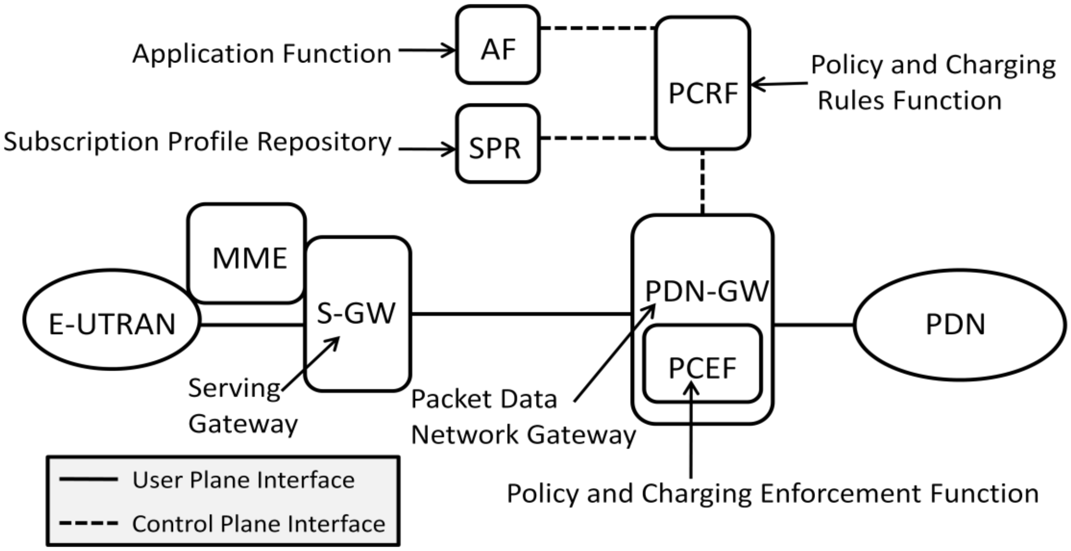

Application Function
AF
PCRF
Policy and Charging Rules Function
Subscription Profile Repository
SPR
MME
E-UTRAN
S-GW
PDN-GW
PCEF
PDN
Serving Gateway
Packet Data Network Gateway
Policy and Charging Enforcement Function
User Plane Interface
Control Plane Interface

Figure 2: The main elements of the LTE PCC framework and their general relationship to each other

QoS Control – Network-Initiated vs. Terminal-Initiated

As discussed by Olsson et al. (2009) in greater detail, there are two different methods available to establish a dedicated bearer for a given level of QoS in a LTE network: network-initiated QoS control and terminal-initiated QoS control. In network-initiated QoS control, the network signals the UE to establish a dedicated bearer with a given level of QoS. Ultimately, it is the PCRF that makes this decision, although it may consult the AF and/or SPR in the process. The exact details of this process depend on a number of factors and are not central to this paper, the key is that with network-initiated QoS control, it is the responsibility of the network to detect and infer what QoS resources are needed by the user or application, without explicitly being told.

In terminal-initiated QoS control, it is the terminal that signals the network and requests that a dedicated bearer with the desired level of QoS be established.[14] This means that the terminal must be aware of the specifics of how QoS is handled in the access network (whereas in

[14] "The purpose of the UE requested bearer resource allocation procedure is for a UE to request an allocation of bearer resources for a traffic flow aggregate. The UE requests a specific QoS demand (QCI) and optionally sends a GBR requirement for a new traffic flow aggregate. If accepted by the network, this procedure invokes a dedicated EPS bearer context activation procedure (see subclause 6.4.2) or an EPS bearer context modification procedure (see subclause 6.4.3)." (3GPP, 2008c).

network-initiated QoS control, terminals can be access-agnostic when it comes to the QoS mechanisms available in a network) and be able to interface with the network to convey the QoS request (typically accomplished using an Application Programming Interface [API]) (Olsson, Sultana, Rommer, Frid, & Mulligan, 2009). However, terminal-initiated QoS control means that a PCRF is not needed to send QoS information to the network (although a PCRF can still be used, if desired, to authorize QoS requests made by terminals) (Olsson, Sultana, Rommer, Frid, & Mulligan, 2009).

Roaming: Home-routed vs. Local-breakout

There are two main roaming scenarios that are supported by the PCC framework in LTE: 'Home Routed' and 'Local Breakout'. In Home Routed roaming, the user in the visited network (i.e. the user who is roaming) is connected to the PDN through a PDN-GW that resides in the home network. Thus, all traffic for that user is routed from the visited network (where the roaming user is connected to the visited E-UTRAN) back through the home network before it exits to external packet networks (e.g. the internet). In Local Breakout roaming, the user in the visited network is connected to the PDN through a PDN-GW in the visited network. Thus, all traffic for that user is routed through the visited network only and never enters the home network.

The PCC architecture was designed to enable the PCRF in the home network (H-PCRF) to communicate with the PCRF in the visited network (V-PCRF) and, when allowed by the visited network, control and authorize all resources for roaming users in the visited network (Olsson, Sultana, Rommer, Frid, & Mulligan, 2009). The exact QoS control the H-PCRF has over its roaming users depends upon which PCC network elements are connected and how they are configured on both networks. In some cases, the V-PCRF may be allowed to either accept or reject (but not change) policy decisions made by the H-PCRF thereby allowing the visited operator some degree of control over the resource usage in its radio access network (i.e. E-UTRAN). In some home routed roaming scenarios, there is no need for a V-PCRF at all. Depending on which roaming scenario is supported (i.e. home routed vs. local breakout), the PCEF that is involved in enforcing policies may be located in either the home or visited network (since the PCEF always resides in the PDN-GW). Where home routed roaming is supported, the PCEF is located in the home network and controlled by the home operator since traffic will be routed through the home PDN-GW.

4. Technical Design Decisions: Choices and Policy Implications

During periods of network congestion when there are insufficient resources for everyone, it is desirable to give some requests for resources preferential treatment. To do so, one must be able to distinguish between different resource requests. There are many different ways public safety may want to differentiate these requests. As we discussed in section 2, in a single service environment (e.g. a voice-only system), distinctions are often made by user type, agency, and role of the user (as is the case in WPS). In a broadband environment, there are even more ways in which one may want to distinguish resource requests, such as by the type of application being used.

While the exact QoS requirements for public safety are outside the scope of this paper, in this section we will identify the main factors that may be involved in differentiating resource requests in a broadband priority access system. For example, in the following table, we propose several likely distinctions that commercial and public safety broadband network operators may make to differentiate resource requests. In the same table, we also identify the some of the possible levels of granularity, and provide specific examples for each distinction.

	Distinction	Granularity	Examples
Commercial	User Identity	Type of Subscription	Business vs. Standard
			High Price Subscription vs. Low Price Subscription
			Post-paid vs. Pre-paid
	Service Type	Real-time vs. Non-real-time application	Voice vs. Web browsing
		Type of real-time application	Voice vs. Video
		Importance of application	911 Voice vs. Other Voice
Public Safety	User Identity	Level of Government	Local vs. State vs. Federal
		Type of Agency	Fire vs. Police vs. EMS
		Rank of User	Officer vs. Chief
	Device Type	Mobility of device	Portable vs. Car-mounted vs. Fixed Location
		User-issued device type	Handheld vs. Laptop
		Fixed device type	Fixed sensor vs. Fixed Video Camera
	Service Type	Real-time vs. Non-real-time application	Voice vs. Web browsing
		Type of real-time application	Voice vs. Video
		Importance of application	Emergency Voice vs. Routine Voice
	Location of Usage	Within a Jurisdiction	Building vs. Highway vs. Rural area
			Government vs. Commercial vs. Residential Building
		Within Jurisdiction vs. Outside Jurisdiction	Local Responder vs. Neighboring Jurisdiction Responder
	Time of Usage	Time of Day	Day vs. Night
			Busy Hour vs. Off-peak
	Situation of Usage	Type of Event	4-Alarm Fire vs. Hurricane vs. Terrorist Attack
	Network Used	Roaming on other Public Safety Network	Network for Region A vs. Network for Region B
		Roaming on Commercial Networks	Network for Carrier A vs. Network for Carrier B

Table 3: Key Factors that may be used to differentiate resource requests on a wireless broadband network, the possible levels of granularity, and specific examples for each distinction

These factors can be static (i.e. the value changes rarely, if ever) or dynamic (i.e. the value can change frequently). If the values for all the factors that are relevant for determining public safety's needs are either known ahead of time (for static factors) or can be detected by the network (for dynamic factors) then it is possible to predefine a set of rules that determines how priority is assigned based on the instantaneous value of these factors (i.e. an automated priority system). In this section, we will focus on how technical design decisions can enable an automated priority system for public safety on a commercial broadband network (while section 5 will focus on operational design decisions that can enable human intervention into priority decisions). To do so, this section will first identify what functionality is provided by a LTE-based network, specifically the mechanisms which enable preferential treatment on the network, and then we study a number of issues that will likely arise in crafting agreements that ensure public safety's needs will be met when using commercial broadband networks.

4.1. What is Possible with Technical Design Decisions

In this section, we focus on the functionality provided by LTE that may be used to meet public safety's needs on a commercial network. We identify the relevant network elements and functions and discuss how they could potentially be used to provide public safety users with the preferential treatment they may need. More specifically, we study the ability of a LTE-based network to differentiate requests based on static factors (e.g. user identity) or dynamic factors (e.g. location) and allocate resources accordingly such that public safety's required QoS characteristics (e.g. bit rate, packet delay, and packet error loss rate) are satisfied. We will provide specific examples of which LTE mechanisms could be used, and also how they could be used to best meet public safety's needs.

In LTE, it is possible to differentiate one user's traffic from another (e.g. a public safety user from a commercial user) and one application's traffic from another (e.g. a video stream from a web browsing session). For example, the bearer mechanisms (discussed in section 3.2) built into LTE enable traffic from applications with similar QoS requirements to be (virtually) separated from traffic with different QoS requirements. In addition, these bearers are established separately for each user (based on subscription and other identity information), thus separating one user's traffic from another's and allowing each to be treated differently. LTE is able to separate this traffic in both the upstream and downstream directions using the TFT mechanisms discussed in section 3.2 and these TFTs enable differentiation and filtering at the packet-level using IP header information.

In LTE, during periods of congestion on the network, it is possible to provide preferential treatment to the traffic from one user or application in a variety of ways. It is possible to block, drop (i.e. preempt) or reduce the QoS of the communications of one user or application. It is possible to ensure packets are prioritized over others such that predetermined QoS characteristics are met, that established sessions are guaranteed a minimum bit rate, and that individual users do not use more than a preset amount of network resources.

For example, the ARP mechanisms discussed in section 3.2 enable two key pieces of functionality: (1) lower priority communication sessions can be preempted by higher priority sessions, and (2) lower priority requests for resources can be blocked if resources are being used for higher priority sessions. Session dropping and/or blocking can be used to give preference to one user over another, or for a given user, give preference to one application over another. For instance, during periods of exceptional congestion (e.g. an emergency incident or a disaster situation) by simply having the eNodeB drop bearers with lower ARP priority level values and preemption vulnerability, additional capacity can be freed up for higher priority bearers.[15] Alternatively, as discussed in (3GPP, 2008b), an operator could map the voice component and the video component of the same video telephony session to separate bearers with different ARP parameters, with the voice bearer receiving a higher ARP priority level than the video bearer. During times of congestion, the eNodeB can then drop the bearer carrying the video component, and not affect the bearer carrying voice at all, allowing at least voice communications to continue during times of congestion.

As another example, the QCI mechanism discussed in section 3.2 ensures that the treatment a packet receives at each node in the network is tied to specific QoS characteristics (e.g. packet delay budget and packet error loss rate) and is subject to varying levels of prioritization. Using QCI values, the required QoS characteristics of each bearer are understood by the network without the need to specify individual QoS characteristics (e.g. packet delay budget) for each bearer.

Additionally, preference can be given by guaranteeing or limiting the bit rate a bearer receives, for example, using the GBR or UE-AMBR mechanisms discussed in section 3.2. The GBR mechanism ensures that accepted sessions are guaranteed their requested bit rate throughout the network (or are dropped subject to their associated ARP parameters). The UE-AMBR mechanism, on the other hand, ensures that the resources used by an individual user (even one with multiple sessions and applications running) never uses more resources than a preset aggregate maximum bit rate.

Finally, in LTE it is possible to roam from one network to another, but still receive preferential treatment on the visited network. For example, in network configurations where roaming QoS is handled using the V-PCRF and ‘local breakout’ mechanisms discussed in section 3.3, it is the roaming network that decides what level of QoS a user roaming on their network receives. However, in ‘home routed’ network configurations using the H-PCRF, also discussed in section 3.3, it is the home network that controls the QoS the roaming user receives, even though the resources allocated are on the visited network.

[15] There is a relationship between the frequency with which sessions are dropped and the frequency with which sessions are blocked from starting; decreasing one increases the other (Peha & Sutivong, 2001). In the work by Bao et al. (2006), it was assumed that it is much worse to preempt a session that is in progress than it is to block that session from even starting in the first place. That paper focused on developing a policy whereby, when an emergency is occurring, resources allocated to commercial users is limited and capacity is reserved exclusively for future public safety use. While this limits the number of commercial sessions that are preempted, it also results in more commercial sessions being blocked initially. However, valuing dropping vs. blocking, and determining the exact implementation of ARP is outside the scope of this paper.

Thus, it is technically possible under the LTE standard to provide several priority-related capabilities that are likely to be important to public safety users. These capabilities mean that during times of congestion, a LTE-based automated priority system with well crafted rules can likely meet the QoS requirements of the individual users or applications public safety deems important, whether or not public safety is roaming on commercial networks or using their own dedicated systems. However, crafting the agreements and designing the rules that will govern an automated priority system face a number of challenges and can have a number of policy implications, as we will discuss in the next section.

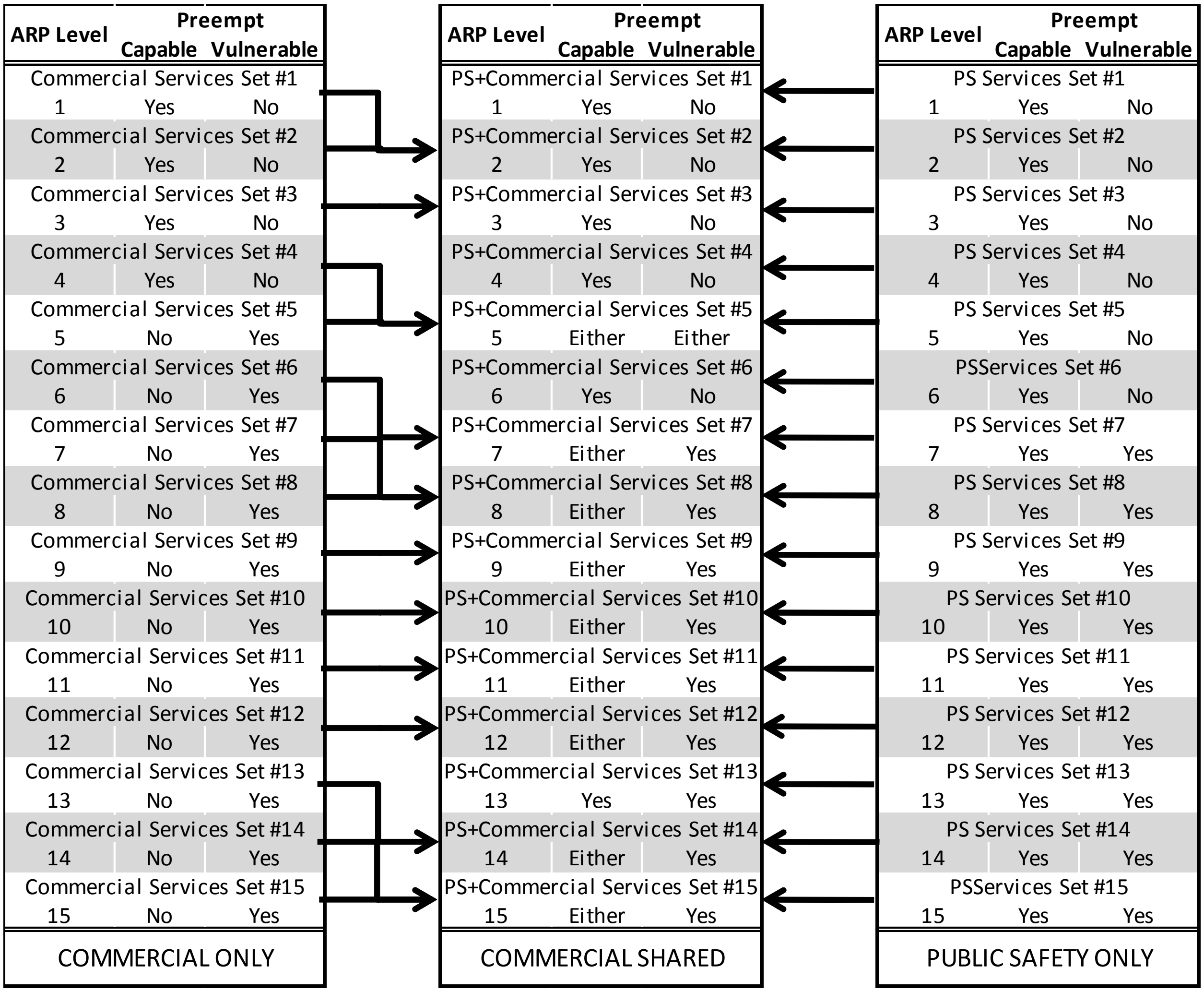
4.2. Policy Implications for Agreements

Public safety use of commercial networks will likely require agreements to be crafted to govern this use. In this section, we will study a number of issues that will likely arise in crafting agreements that ensure public safety's needs will be met, including: the minimum level of specification required to ensure commonality across agreements, how agreements should handle the potential need for changes to technology standards or even new products to be developed if in the future public safety's needs diverge from those of commercial users, and how agreements should handle the possibility of vendor and/or operator-specific implementations even when standards are used.

Technical Requirements in Agreements

To facilitate roaming, particularly roaming with preferential treatment, agreements (e.g. Service Level Agreements or SLAs) may need to be negotiated between public safety agencies and commercial operators. These agreements may be necessary since there are a number of technical design decisions that must be settled to ensure that the QoS public safety users experience is consistent, or at least sufficient, across networks. For example, not all network elements are required for an LTE deployment and not all the features of LTE are supported by every network configuration. Indeed, for some of the network features to be supported, coordination will be required between both sides. For example, as discussed in (Motorola, 2010), the PCC framework in LTE and the related network elements like the PCRF are optional functions of the LTE architecture. If, for instance, public safety needs direct control over the QoS its users receive when roaming, then, in order to provide this functionality, public safety will have to include a PCRF in its network and the commercial network will have to support and allow home routed traffic for users roaming on its network (as discussed in section 3.3).

As another example, the QCI and ARP parameter values used and the manner in which the static and dynamic factors (discussed earlier in this section) are mapped to these values may need to be agreed upon by both the commercial operators and public safety agencies. Agreeing to apply the QCI and ARP parameters consistently across both commercial and public safety networks will ensure public safety users receive the QoS they require, even while roaming. For example, of the 15 currently defined ARP priority levels, the LTE standard (3GPP, 2008d) indicates that values 9 to 15 (i.e. the lowest 7 priority levels) should be used for roaming UEs. If the lowest 7 ARP levels are insufficient to meet public safety's needs when roaming on commercial networks, then public safety may find it useful to agree with commercial operators on a consistent policy for the ARP values public safety has access to when roaming. As discussed in (Motorola, 2010), the challenge with settling on ARP priority levels in agreements will be determining how to map both commercial users and public safety users (and devices, applications, and other key factors) onto the same set of ARP values. For instance, does the value assigned to a public safety user change as they roam onto a commercial network, or do their values remain constant regardless of what network they are using? Do some commercial users receive higher priority than public safety users and, if so, which users? While we will not attempt to map the key public safety factors discussed earlier in this section to ARP values here, as it is outside the scope of this paper, the following figure illustrates a hypothetical mapping of commercial services and public safety services onto a single consistent ARP policy.



ARP Level	Preempt	
	Capable	Vulnerable
Commercial Services Set #1		
1	Yes	No
Commercial Services Set #2		
2	Yes	No
Commercial Services Set #3		
3	Yes	No
Commercial Services Set #4		
4	Yes	No
Commercial Services Set #5		
5	No	Yes
Commercial Services Set #6		
6	No	Yes
Commercial Services Set #7		
7	No	Yes
Commercial Services Set #8		
8	No	Yes
Commercial Services Set #9		
9	No	Yes
Commercial Services Set #10		
10	No	Yes
Commercial Services Set #11		
11	No	Yes
Commercial Services Set #12		
12	No	Yes
Commercial Services Set #13		
13	No	Yes
Commercial Services Set #14		
14	No	Yes
Commercial Services Set #15		
15	No	Yes
COMMERCIAL ONLY		

ARP Level	Preempt	
	Capable	Vulnerable
PS+Commercial Services Set #1		
1	Yes	No
PS+Commercial Services Set #2		
2	Yes	No
PS+Commercial Services Set #3		
3	Yes	No
PS+Commercial Services Set #4		
4	Yes	No
PS+Commercial Services Set #5		
5	Either	Either
PS+Commercial Services Set #6		
6	Yes	No
PS+Commercial Services Set #7		
7	Either	Yes
PS+Commercial Services Set #8		
8	Either	Yes
PS+Commercial Services Set #9		
9	Either	Yes
PS+Commercial Services Set #10		
10	Either	Yes
PS+Commercial Services Set #11		
11	Either	Yes
PS+Commercial Services Set #12		
12	Either	Yes
PS+Commercial Services Set #13		
13	Yes	Yes
PS+Commercial Services Set #14		
14	Either	Yes
PS+Commercial Services Set #15		
15	Either	Yes
COMMERCIAL SHARED		

ARP Level	Preempt	
	Capable	Vulnerable
PS Services Set #1		
1	Yes	No
PS Services Set #2		
2	Yes	No
PS Services Set #3		
3	Yes	No
PS Services Set #4		
4	Yes	No
PS Services Set #5		
5	Yes	No
PSServices Set #6		
6	Yes	No
PS Services Set #7		
7	Yes	Yes
PS Services Set #8		
8	Yes	Yes
PS Services Set #9		
9	Yes	Yes
PS Services Set #10		
10	Yes	Yes
PS Services Set #11		
11	Yes	Yes
PS Services Set #12		
12	Yes	Yes
PS Services Set #13		
13	Yes	Yes
PS Services Set #14		
14	Yes	Yes
PSServices Set #15		
15	Yes	Yes
PUBLIC SAFETY ONLY		

Figure 3: Hypothetical mapping of commercial services and public safety services onto a single consistent ARP policy

Commonality across Agreements

Some degree of commonality across agreements based on meaningful guidelines will likely reduce the need for thousands of different agreements to be negotiated separately, while still allowing for divergence between agreements. One challenge with creating commonality is that there is no single public safety entity to establish meaningful guidelines, but instead there are many separate agencies (e.g. in the US there are more than 50,000 local, state, and federal public safety agencies (Hallahan & Peha, 2008; 2009; 2010)). Thus, by identifying an entity (or possibly creating one) whose job it is to solicit input from the many public safety agencies and establish appropriate guidelines for agreements, the burden placed on each of these individual agencies will likely be reduced.

For example, in the US it is anticipated that there will be many regional public safety networks and potentially several commercial networks onto which a given public safety user can roam. If bilateral agreements are pursued between all of these entities it could lead to hundreds or even thousands of separate agreements which may prove onerous. Multilateral agreements or bilateral agreements based on established guidelines can provide increased commonality (which reduces transaction costs) while still allowing for some degree of divergence between agreements (which promotes agreements that serve regional needs). The more specific the agreement, the more likely the agreement can ensure the needs of public safety will be met. However, overly specific agreements can have a stifling effect on innovation over time and diversity of product and service offerings across providers. Thus, there is a tradeoff between the two, and the goal is to identify the minimum level of specificity that can meet public safety’s needs, while enabling innovation and evolution to occur.

For example, key components such as the air interface (e.g. LTE in the US) used by devices to communicate over the radio channel with cell sites, will likely need to be specified in any agreements to ensure both public safety devices and commercial devices can communicate with the same cell sites. If the air interface is not common to all agreements, in order to ensure interoperability, each agency would have to negotiate agreements that specify which air interface will be used, not just with every commercial network, but with every regional public safety network that they may roam on. However, it is not necessary for the air interface (or many other components for that matter) to be specified statically; instead, the agreement can indicate that network upgrades are allowed, as long as upgrades provide backwards compatibility or upgrade cycles are coordinated such that both public safety and commercial users move to the new technologies in some agreed upon timeframe. In addition, it may not even be necessary to require a specific standard be used for the entire network, instead a family of standards may be sufficient as long as they are backwards compatible (e.g. in the US, agreements could likely state that either 3GPP Release 8 or 9 equipment is sufficient since both are compatible to some degree).

Agreements for Vendor/Operator-Specific Implementations

Agreements should be structured such that they ensure public safety's needs are met even when the technological standards on which the agreements are based allow for vendor and/or operator-specific implementations. That is, vendors and/or operators may make different decisions in their implementations of an LTE network that could lead to a different experience on different networks, even when these networks are based on the same technological standard. For example, in LTE, the manner in which a node handles packet forwarding functions can be implementation-specific to some degree.

More specifically, in LTE, packet forwarding functions[16] (e.g. the scheduling algorithms used at a node) use standardized QoS parameters as inputs, but the algorithms themselves are not standardized. Instead, operators and vendors are allowed to make their own decisions as to how to implement these functions (Olsson, Sultana, Rommer, Frid, & Mulligan, 2009). For example, if a scheduler always favors every higher priority packet over lower priority packets,[17] during times of congestion, this scheduler may consistently starve lower priority bearers of any bandwidth. Thus, an operator may design their scheduler such that some minimum level of bandwidth is always reserved to be allocated to the lowest priority bearers. This would ensure the low priority bearers never completely starve, even if some higher priority packets are delayed and/or dropped as a result. Operators on different networks could make different decisions when implementing this specific functionality and this could potentially lead to the traffic from public safety users (even with standardized QoS parameters attached) experiencing different levels of performance in different networks, even under identical conditions. In this paper, we make no judgment as to the likelihood of this particular problem; instead this is simply a specific example of how an implementation-specific problem could arise.

The options for crafting agreements to handle vendor/operator-specific implementations are either to make the agreements independent of those vendor/operator specific decisions (e.g. by specifying performance requirements and allowing vendors and operators to make their own decisions as long as performance guarantees are met) or to include in the agreements breakout pieces that provide rules to govern each vendor. In either case, it is useful to have some entity that has sufficient expertise in these technical issues (i.e. expertise with the technical standards

[16] The specifics of many possible techniques for the packet forwarding treatment used in the radio access network are discussed in more detail in (Dahlman, Parkvall, Sköld, & Beming, 2008; Sesia, Toufik, & Baker, 2009), but are outside the scope of this particular paper. In addition, packet forwarding treatment within the EPC is typically handled using existing IP-based techniques like Differentiated Services (DiffServ) by mapping the QCI parameters of the EPS bearer to the DiffServ Code Point (Olsson, Sultana, Rommer, Frid, & Mulligan, 2009). We will not go into the IP-based techniques used in the EPC as they are also outside the scope of this paper.

[17] For example, as indicated in (3GPP, 2008a), scheduling between different bearers will be primarily done based on the packet delay budget associated with the QCI of that bearer. If the target packet delay budget cannot be met for some of the bearers (across all UEs that have sufficient radio channel quality) then the priority of the QCI associated with each bearer shall be taken into account when scheduling. This happens as follows: the scheduler will meet the packet delay budget for all bearers of priority level N in preference to meeting the packet delay budgets of bearers with priority level N+1.

and vendor implementations, for instance, through demonstration or independent testing), but also a deep understanding of public safety's needs, and the responsibility to advise public safety with the necessary decisions and provide guidelines to reduce the burden on individual agencies.

Agreements Now for Needs that May Change Over Time

Agreements must be constructed now using current technologies and standards, and defining perfectly everything needed in year 10 may be very difficult to do today. Even if this is possible, over time, the needs of both commercial and public safety users may change, and possibly do so in diverging directions. Regardless, the agreements will need to guarantee public safety's needs are met, especially as technology standards and products evolve to meet their changing needs. In particular, QoS mechanisms such as the QCI, ARP, and MBR parameters (which may be of particular interest to public safety) are available and standardized now, but in the future it may be useful to change how these parameters are defined. The likelihood of changes being required for these particular parameters is outside the scope of this paper. Instead, these parameters are presented as specific examples of how network mechanisms and parameters that are standardized and agreed upon initially (such that they meet public safety's needs today), may not always meet public safety's needs in the future.

For example, there are currently 15 values of ARP priority level defined in the LTE standard. Even if 15 values are sufficient now to differentiate all public safety and commercial traffic, there is no guarantee 15 ARP priority levels will always be sufficient. If, at some point, 15 levels are no longer sufficient to meet public safety's needs then there is value in defining new, public safety-specific levels. Similarly, 9 QCI values are currently defined in the LTE standard and, like ARP values, even if 9 values are sufficient now, in the future there is no guarantee that the 9 currently defined QCI values will always meet public safety's requirements. Unlike ARP values, QCI values are mapped to specific QoS characteristics (e.g. packet delay budget and packet error loss rate), as discussed in section 3.2, and thus additions and modification over time may be useful as new applications evolve which have QoS requirements that differ from the QoS requirements of today's applications.

Furthermore, it may be useful to define additional QCI values in order to provide additional levels of prioritization within the same set of QoS characteristics. As discussed in section 3.2, different applications mapped to the same QCI value (on the grounds of QoS characteristics alone), will have the same priority level. In times of congestion, the nodes in the network will treat all packets with the same QCI value the same, even if the relative importance of the two is different. As an example, QCI values 8 and 9 are identical to one another in terms of the standardized QoS characteristics they map to (i.e. the packet delay budget for both is 300ms and the packet error loss rate for both is 10^{-6}); however, the priority values associated with QCI values 8 and 9 are different (i.e. QCI 8 has priority level 8 and QCI 9 has priority level 9). Thus, while nodes will generally treat the packets from bearers with these two QCI's the same, during times of congestion, if the QoS characteristics for both cannot be met the packets associated with the higher priority level will receive preference. So even if the current set of 9 QCIs provide the necessary QoS characteristics to meet public safety needs, it could turn out to be useful to define additional QCIs to provide a finer level of granularity of priority for each set of QoS characteristics.

Another example is the MBR parameter which specifies the maximum bit rate that a GBR bearer is allowed. As discussed by Olsson et al. (2009), the MBR parameter along with the GBR parameter (which guarantees a minimum bit rate for the bearer) could be used to support rate-adaptive codecs (wherein the network provides a minimum bit rate using the GBR parameter and, when possible, a higher bit rate using the MBR parameter). For example, this could enable video applications which provide a higher resolution when the bandwidth is available, but can be throttled back to a lower resolution when congestion occurs. Given the role video applications are expected to play in public safety broadband networks, this type of functionality could become particularly useful for public safety. However, in the current standard (3GPP, 2008b) the MBR parameter of a particular bearer has to be set equal to the GBR value for that bearer. This makes it impossible to use the MBR mechanism in its current form to throttle back applications. Instead, GBR bearers can only be disconnected completely if the eNodeB can no longer support the bit rate they require.[18] In the future, public safety may desire this sort of functionality, which would likely necessitate updates to the current standards.

While there are many different ways one can imagine public safety's needs will evolve in the future, it is impossible to know for sure what will occur. If public safety's needs evolve in the same way as commercial needs (such that any changes to products and standards required by one are the same as those required by the other) it is less of a challenge to address these changes. However, even if public safety's needs diverge from those of commercial users, not all required changes are equal. Some changes require modifying or adding to already standardized parameters and mechanisms, while other changes may require all new functionality be added. This can affect the cost to standardize the change, and the impact this change has on affected networks.

Modifications to parameters already specified may mean that the modifications can be accomplished relatively easily and the cost to implement the change is small (although any changes will require someone to understand the needs of the public safety community and act on their behalf). If additional modifications are required, likely candidates are the QCI and ARP parameters, as they are of particular interest to public safety when ensuring their QoS needs are

[18] The EPC does not support an eNodeB-initiated modification procedure for GBR bearers. If, for some reason, an eNodeB can no longer sustain the bit rate required for a GBR bearer that has already been set up by the EPC and is currently active, then the eNodeB can only trigger a deactivation of that bearer. So, for example, if a user is authorized for a GBR to support a real-time video application and that user moves to the edge of the cell where the radio conditions are worse, and the eNodeB is no longer able to sustain the required bit rate, it can only drop the bearer as there is no mechanism for the eNodeB to provide feedback to the EPC to modify the GBR of the bearer (3GPP, 2008b).

met. In this case, the modifications are relatively easy in part because (1) flexibility for additional values was already designed into LTE (i.e. there are already extra bits available for manipulation and standardization), (2) modifications can be made without affecting parameter values that were already defined (i.e. standardizing QCI value 10, does not affect existing QCI value 9) and therefore without affecting all legacy devices that use only those values, and (3) modifications to components within that infrastructure to accommodate devices that use new QCI and ARP values can likely be handled through software updates rather than hardware replacements.

As specific examples, it may be possible to define additional QCI and ARP values with limited changes to the technical specifications of the standard. More specifically, the QCI value is carried in octet 3 of the EPS quality of service information element, and with 8 bits per octet, this octet can take up to 2^8, or 256, different values (3GPP, 2008c). Currently, only 9 of the 256 values are standardized, the rest of the values are either reserved for future use or operator-specific QCI's. This implies that if standardized, LTE could support hundreds of QCI levels without major revisions to the rest of the specifications (and the costs associated with major revisions). In addition, the ARP parameter is contained in octet 5 of the Bearer Quality of Service field, and this octet uses two of the bits for the preemption vulnerability and preemption capability flags and four of the bits to represent the 15 ARP priority levels that are currently standardized (3GPP, 2008c). The remaining two bits in the octet are currently designated as spares. This implies that it may be possible to increase the number of ARP priority levels available by simply making use of the two spare bits (theoretically increasing the number of priority levels available from 16 to 64).

While some changes to the standard can be handled by software upgrades, it is possible to imagine others that require major physical changes to legacy equipment, and in the worst case, legacy equipment owned by commercial carriers with no incentive to upgrade (and thus this cost may be passed to public safety). Even when changes can be handled with software updates, the costs may differ depending on if software upgrades are required for the network equipment, handsets, or both. For example, when a node receives a QCI value it doesn't understand, it will simply choose a QCI value that it understands instead (3GPP, 2008c), potentially leading to an inconsistent QoS experience. In this example, changes made to the standard QCI values may be handled through software updates to the affected network nodes. However, in those networks with terminal-initiated QoS control (discussed in section 3.3), changes to standardized QoS values could make it more important to update handsets that have already been deployed (and potentially their vendor-specific APIs as discussed by Olsson et al. (2009)). This task is complicated by the fact that these devices are not all under the direct control of the network operator.

5. Operational Design Decisions: Automated versus Human Intervention

In section 4, we studied technical design decisions that will enable preferential treatment of public safety traffic on commercial networks, assuming that public safety's QoS requirements can be satisfied by a fully automated priority system. However, it is possible that, in some cases, allowing humans to intervene and affect the preferential treatment specific individuals or agencies or applications receive results in a noticeable improvement in the QoS public safety experiences, particularly during periods of congestion. In this section, we will discuss what can be accomplished with an entirely automated priority system, and identify situations where human intervention may be beneficial.

In an automated priority system, the priority parameters (in LTE examples are the QCI and ARP values) are assigned according to policies and decision rules that are based ONLY on factors that the system either (1) knows the value of *a priori* or (2) can detect the value of without any human intervention. An automated priority system's decisions can be based on static factors and/or dynamic factors that the network is aware of and can detect changes in. Static factors have a fixed value that rarely, if ever, changes and are typically known *a priori* by the network (e.g. user identity, such as police officer, could be stored by the network in some sort of subscriber profile so it is automatically known by the network). Dynamic factors do not have a fixed value, but instead the value they hold can change frequently and therefore must be measured by the network. For example, 'Location' could be a dynamic factor that takes as its value the current location of the device and, for instance, could be used to prioritize devices based on their proximity to designated locations. In order for priority decisions to be based on the desired factors, the potential values of each static and dynamic factor are mapped to some predefined decision rule (e.g. if roaming = yes, priority level is X; if roaming = no, priority level is Y; where, in LTE for example, X could equal ARP priority level 2 and Y could equal ARP priority level 1).

With a purely automated priority system, a network operator can provide the necessary preferential treatment in a great many circumstances and, in many cases, its ability to prioritize resource requests adequately is only limited by what factors have or have not been mapped to priority levels. Indeed, the decision rules in an automated priority system could be anywhere from very simple to extremely complex; it is up to the network implementation and the agreements reached between commercial operators and public safety. A simpler set of rules may have the advantage of easier implementation, while complex rules that make use of many different factors may do a better job of differentiating users (which may be beneficial to public safety). But, there is some limit to the added benefit provided from making decision rules even more complex. In fact, in some cases, an automated priority system may not be able to provide the desired functionality regardless of how sophisticated the decision rules and mappings of priority levels are.

In some cases, the ideal resource allocation depends on some factor that is (1) known to the public safety user or other individual responsible for that device and (2) unknown and undetectable by the network on its own. Unlike factors such as application type or roaming status, factors like the intention of a first responder or their perceived level of danger in the current incident cannot be measured or detected by the network and therefore cannot be used for resource allocation using an automated priority system. Put another way, an automated priority system could base priority decisions on the fact that a police officer is making a voice call from his patrol vehicle on a highway outside of his normal jurisdiction; it just can't distinguish whether he is pursuing a known fugitive or driving back to the police station, nor can it make a priority assignment based on that distinction.

To handle these cases, a non-fully-automated priority system could be designed such that the decision rules that dictate who or what receives which preferential treatment depend, at least in part, on factors that require human intervention to set their current state. This intervention-enabled priority system could still provide all of the functionality that an automated priority system provides, with human intervention as an added capability. In the previous example, this means that the police officer could intervene with an intervention-enabled priority system and (by some mechanism) indicate that this particular session is of elevated importance because he is pursuing a suspect. But if he doesn't intervene, the system can still make priority assignments based on the other factors it knows (for instance, user identity, application, and roaming status), just as a fully automated priority system can.

While allowing human intervention can enable a system to make better decisions about prioritization and resource allocation in some situations, human intervention also brings some complications that must be addressed. Designing a system that uses human intervention as an input necessarily requires that humans be available with both the authority and expertise to make decisions that affect the network, and do so within the appropriate time constraints. There can be first responders from many different public agencies simultaneously responding to multiple emergencies, each with its own incident commander. Even without roaming, it must be determined who is allowed to intervene in ways that affect who is allocated resources, and how. With the addition of priority roaming, solutions must be found that are effective for all first responders that are roaming onto commercial networks, while also treating commercial traffic appropriately. Moreover, someone must define methods by which those with situational knowledge (whether they are incident commanders, dispatchers, or individual first responders) communicate that knowledge back to the networks, and those methods would be common across the many public safety agencies and commercial networks. It is also possible that the additional technical and operational complexity will also affect cost.

An intervention-enabled priority system would be desirable if the incremental benefit it provides over an automated priority system is greater than the associated incremental costs of implementing such a system. Thus, decision makers should understand the needs of the public safety community, the functionality provided by automated priority systems, and the additional functionality provided by intervention-enabled systems and then balance these against the additional complexity and challenges human intervention presents before deciding on which method to employ. Here we will present a few potential operational arrangements for enabling human intervention as well as a few examples of possible technical implementations, while future work may study additional decisions and their tradeoffs in greater detail.

LTE makes possible a wide range of options for enabling human intervention in priority decisions. Each option has advantages and disadvantages and which one is best is outside the scope of this particular paper, but we will present three possible options as examples. One option may be to give a centralized public safety entity full control over the QoS decisions made for their users, even while those users are roaming and therefore these decisions affect commercial users as well. Another option may be to allow commercial operators to hold final control over the QoS decisions that affect their network, while public safety provides input to the commercial operator on these decisions. A third option is to leave the decision up to individual public safety users, and allow them to affect the QoS they receive in response to their current situation.

Here are three technical examples (which were discussed in more detail in section 3.3) of how to implement these options for human intervention; other approaches are also possible. For example, by making use of the (optional) PCRF function and 'home-routed traffic' when public safety users are roaming, the QoS level these users receive will be controlled by the H-PCRF. Therefore, if a public safety representative maintains control over the H-PCRF (or multiple H-PCRFs if there are multiple regional public safety networks), then public safety could potentially intervene and update the policies to reflect current situations, even for users who are roaming on commercial networks with no action required by the commercial operator. Or, for example, if the PCRF function and 'local breakout' are employed, a roaming public safety users' QoS would be controlled by the V-PCRF in the roaming network. Thus, if a commercial operator maintains control over their V-PCRF, they can control who intervenes to update the policies and subscriber privileges the network uses to make QoS decisions.   The operator may receive public safety input in a variety of ways, but the operator has responsibility for the network element that controls priority. As another example, by making use of terminal-initiated QoS control, public safety users may be able to provide the necessary feedback (e.g. by pressing an emergency button when in distress) which informs the network of the current state of their needs (e.g. by requesting a bearer of desired QoS level be established). In this option, the control over human intervention is passed on to the first responders in the field, who can make decisions based on their instantaneous situational needs.

6. Conclusions

Wireless broadband functionality could revolutionize the way public safety responds to emergencies by bringing capabilities to first responders they have never before had access to. To bring this functionality to their users, both public safety and commercial operators will likely deploy wireless broadband networks. Providing public safety users roaming access to commercial networks, on a priority basis, can provide greater aggregate capacity, geographic coverage, and service reliability than would be available from the dedicated public safety networks alone. This paper studied technical and operational issues associated with providing public safety users with priority roaming access assuming there are several broadband networks run by both public safety and commercial operators, where both commercial and public safety users compete for resources during periods of network congestion, and use many different applications which can have different bit rate and QoS requirements.

This paper has shown that it is technically possible under the LTE standard to provide several priority-related capabilities that are likely to be important to public safety users. Thus, during times of congestion, a LTE-based automated priority system with well crafted rules can likely meet the QoS requirements of the individual users or applications that public safety deems most important, whether or not public safety users are roaming on commercial networks or using their own dedicated systems. That is, using the bearer mechanisms in LTE, it is possible to differentiate one user's or application's traffic from another. It is possible to block or drop the communications of one user or application (e.g. using the ARP parameter) and ensure packets are prioritized over others such that predetermined QoS characteristics are met (e.g. using the QCI parameter). It is also possible to guarantee sessions a minimum bit rate (e.g. using the GBR parameter) and ensure that individual users do not use more than a preset amount of network resources (e.g. using the MBR and AMBR parameters).

This paper also identified possible issues that may arise in crafting the agreements and designing the rules that will govern an automated priority system for public safety users roaming on commercial networks. Agreements may be necessary to ensure that the QoS public safety users experience is sufficient while roaming. For example, some features and network elements are optional in LTE, and to ensure some of the network features that public safety may desire are supported, coordination may be necessary between commercial and public safety networks (e.g. for 'home routed' roaming traffic, the public safety network must implement the optional PCRF element, and the commercial operator must allow this element to interact with their network). Additionally, agreeing to apply the QCI and ARP parameters consistently across both commercial and public safety networks will ensure public safety users receive the QoS they expect, even while roaming.

Additionally, some degree of commonality across agreements based on meaningful guidelines can reduce the need for thousands of different agreements to be negotiated separately, while still allowing for divergence between agreements. That is, if there are many regional public safety networks and several commercial networks deployed and bilateral roaming agreements are pursued, establishing thousands of separate agreements may prove onerous. Agreements based on established guidelines can reduce transaction costs by providing increased commonality while still promoting agreements that serve regional needs by allowing for some degree of divergence between agreements. The goal should be to identify the minimum level of specificity that can meet public safety's needs, while enabling innovation and evolution to occur.

Since technology standards that allow for vendor/operator specific implementations introduce the possibility that different networks using the same standard may provide a different user experience, agreements should be structured so that public safety's needs are met even when network designs can be implementation-specific. For example, in LTE, packet forwarding functions (such as the scheduling algorithms used at a node) use standardized QoS parameters as inputs, but the algorithms themselves are not standardized. This means operators and/or vendors can make their own design decisions which could lead to the traffic from public safety users (even with standardized QoS parameters attached) experiencing different levels of performance in different networks, even under identical conditions. To handle this scenario, agreements can be crafted so they are independent of those vendor/operator specific decisions (e.g. by specifying performance requirements) or to include in the agreements specific sections that provide rules to govern each vendor.

Finally, this paper discussed the challenges associated with crafting agreements now using current technologies and standards and ensuring that public safety's needs are met even though, over time the needs of both commercial and public safety users may change and possibly diverge in some respects. For example, 9 QCI values are currently defined in the LTE standard and even if they are sufficient now, in the future there is no guarantee that they will always meet public safety's needs. Thus, additions and modification to the standardized QCI values over time may be useful as new applications evolve which have QoS requirements that differ from the QoS requirements of today's applications. But not all changes that may need to be made are equal; some are more complex and costly than others. For instance, modifying the QCI parameter can be accomplished relatively easily in part because (1) flexibility to define additional values was designed into LTE (e.g. only 9 of the 256 QCI values that can be set in the EPS quality of service information element are standardized), (2) modifications can be made without affecting parameter values that were already defined or affecting legacy equipment that uses only those parameters (e.g. standardizing QCI value 10, does not affect the existing QCI value 9), and (3) modifications to components within that infrastructure to accommodate devices that use new QCI values can likely be handled through software updates without the need for physical hardware replacements.

In order to reduce the burden placed on individual public safety agencies when establishing guidelines to promote commonality across agreements, or when addressing vendor/operator specific implementations in standards, or when modifying standards to meet public safety's evolving needs, it is useful to identify (or establish) some entity to advise and speak for public safety. This entity, whoever is chosen, should have sufficient expertise in the technical issues, a deep understanding of public safety's needs, the responsibility to solicit feedback from the public safety community, and the authority to act on their behalf.

This paper also demonstrated the technical difference between what can be accomplished with a fully-automated priority system and what can be accomplished with a non-fully-automated, intervention-enabled priority system. In an automated priority system, the priority parameters are assigned according to policies and decision rules that are based ONLY on factors that the system either (1) knows the value of *a priori* or (2) can detect the value of without any human intervention. But, in some cases, the ideal resource allocation depends on some factor that is (1) known to the public safety user or other individual responsible for that device and (2) unknown and undetectable by the network on its own. To handle these cases, a non-fully-automated priority system could be designed such that the decision rules that dictate who receives what preferential treatments depends, at least in part, on factors that require human intervention to set their current state.

This paper has shown that with a purely automated priority system, a network operator can provide the most effective form of preferential treatment in a great many circumstances, but not all. This paper has shown that the intervention-enabled priority system could still provide all of the functionality that an automated priority system provides, but with human intervention as an added capability. For example, an automated priority system can base priority decisions on the fact that a fire fighter is streaming video and is located within his fire department's jurisdiction, whereas an intervention-enabled priority system could also factor in whether or not the fire fighter is in a burning building at that moment.

In this paper, we discussed how LTE makes possible a wide range of options for enabling human intervention in priority decisions and we presented three possible options as examples. One option is to give a centralized public safety entity full control over the QoS decisions made for all their users, users roaming on commercial networks included. Another is to allow commercial operators to hold final control over the QoS decisions that affect their network, with input on these decisions provided by public safety. And finally, the third is to allow first responders to directly affect the QoS they receive based on their current situation. While other approaches may be possible, this paper presented three possible examples for implementing human intervention in a LTE-based network. By making use of the PCRF function and 'home-routed traffic' when public safety users are roaming, the QoS level these users receive can be controlled by a centralized public safety entity through their control of the H-PCRF element located in their network. Or, if the PCRF function and 'local breakout' are employed, a roaming public safety user's QoS could be controlled by the commercial operator through their control of the V-PCRF in the roaming network. Finally, by making use of terminal-initiated QoS control, public safety users may be able to control the QoS level they receive by providing feedback to inform the network of the current state of their needs.

Finally, this paper discussed the fact that while the introduction of human intervention can bring added functionality, it can also introduce additional complexities that must be dealt with. An intervention-enabled priority system would be desirable if the incremental benefit it provides over an automated priority system is greater than the associated incremental costs of implementing such a system. Therefore decision makers should understand the needs of the public safety community, the functionality provided by both fully-automated and intervention-enabled priority systems, and the costs and challenges associated with each before deciding which method to employ.

7. References

3G Americas. (2010, July). *Global 3G status UMTS / UMTS-HSPA/ HSPA+ / LTE.* Retrieved from 〈 http://www.3gamericas.org/documents/Global%20Status%20Update%20July%2027%202010.pdf 〉

3GPP. (2010). *3GPP - Specifications*. Retrieved July 22, 2010, from 〈 http://www.3gpp.org/specifications 〉

3GPP. (2008a, December). *Technical Specification 23.203, 'Policy and charging control architecture (Release 8)'.* Retrieved from 〈 http://www.3gpp.org/ftp/Specs/html-info/23203.htm 〉

3GPP. (2008b, December). *Technical Specification 23.401, 'General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access (Release 8)'*. Retrieved from 〈 http://www.3gpp.org/ftp/Specs/html-info/23401.htm 〉

3GPP. (2008c, December). *Technical Specification 24.301, 'Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3 (Release 8)'.* Retrieved from 〈 http://www.3gpp.org/ftp/Specs/html-info/24301.htm 〉

3GPP. (2008d, December). *Technical Specification 29.212, 'Policy and charging control over Gx reference point (Release 8)'*. Retrieved from 〈 http://www.3gpp.org/ftp/Specs/html-info/29212.htm 〉

Ackerman, R. K. (2003, March). Cellular priority system begins operation. *SIGNAL Magazine* (http://www.afcea.org/signal/articles/templates/SIGNAL_Article_Template.asp?articleid=273).

Bao, J. Q., Guo, L., & Lee, W. C. (2006). Policy-based resource allocation in a wireless public safety network for incident scene management. *Proceedings of the 2006 International Conference on Mobile Computing and Networking, Workshop on dependability issues in wireless ad hoc networks and sensor networks* (pp. 29-34). New York: ACM doi:10.1145/1160972.1160978 .

Chambers, M. D., & Riley, D. H. (2004). Implementing wireless priority service for CDMA networks. *Bell Labs Technical Journal , 9* (2), 23–36, doi:10.1002/bltj.20023 .

Dahlman, E., Parkvall, S., Sköld, J., & Beming, P. (2008). *3G evolution: HSPA and LTE for mobile broadband* (2nd ed.). Oxford: Academic Press.

Federal Communications Commission [FCC]. (2010a, March 16). *Connecting America: The national broadband plan.* Retrieved from 〈 http://www.broadband.gov/plan/ 〉

Federal Communications Commission [FCC]. (1998, September 29). *First report and order and third notice of proposed rulemaking in the matter of the development of operational, technical and spectrum requirements for meeting federal, state and local public safety agency communication requirements through the year 2010.* Retrieved from WT Docket No. 96-86: 〈 http://wps.ncs.gov/documents/fcc98191.pdf 〉

Federal Communications Commission [FCC]. (2010b, May 12). *Order in the matter of requests for waiver of various petitioners to allow the establishment of 700 MHz interoperable public safety wireless broadband networks.* Retrieved from PS Docket No. 06-229: 〈 http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-10-79A1.pdf 〉

Federal Communications Commission [FCC]. (2000, July 13). *Second report and order in the matter of the development of operational, technical and spectrum requirements for meeting federal, state and local public safety agency communication requirements through the year 2010.* Retrieved from WT Docket No. 96-86: 〈 http://wps.ncs.gov/documents/242.pdf 〉

Hallahan, R., & Peha, J. M. (2008). *Quantifying the costs of a nationwide broadband public safety wireless network.* Proceedings of the 36th Telecommunications Policy Research Conference, 〈 http://www.ece.cmu.edu/~peha/costs_of_public_safety_network.pdf 〉.

Hallahan, R., & Peha, J. M. (2010). Quantifying the costs of a nationwide public safety wireless network. *Telecommunications Policy , 34* (4), 200-220, doi:10.1016/j.telpol.2010.01.002 .

Hallahan, R., & Peha, J. M. (2009). *The business case of a nationwide wireless network that serves both public safety and commercial subscribers.* Proceedings of the 37th Telecommunications Policy Research Conference, 〈 http://www.ece.cmu.edu/~peha/profitability_of_public_safety_network.pdf 〉.

Johnson, C. (2010). *Long Term Evolution IN BULLETS.* Northhampton, UK: CreateSpace.

Motorola. (2010, July 19). *Comment before the Federal Communications Commission in the matter of Public Safety and Homeland Security Bureau seeks comment on interoperability, out of band emissions, and equipment certification for 700MHz public safety broadband networks.* Retrieved from PS Docket No. 06-229: 〈 http://fjallfoss.fcc.gov/ecfs/document/view?id=7020549858 〉

National Communications System. (2009). *Fiscal year 2009 report.* Retrieved from 〈 http://www.ncs.gov/library/reports/ncs_fy2009.pdf 〉

National Communications System. (1995, October 19). *Petition for Rulemaking before the Federal Communications Commission in the matter of cellular priority access for national security and emergency preparedness telecommunications.* Retrieved from WT Docket No. 96-86: 〈 http://ecfsdocs.fcc.gov/filings/1995/10/19/151491.html 〉

National Communications System. (n.d.). *Wireless Priority Service (WPS).* Retrieved July 22, 2010, from 〈 http://wps.ncs.gov/ 〉

Olsson, M., Sultana, S., Rommer, S., Frid, L., & Mulligan, C. (2009). *SAE and the Evolved Packet Core: Driving the mobile broadband revolution* (1st ed.). Oxford: Academic Press.

Peha, J. M. (2008, May 26). *A 'successful' policy for public safety communications, Comment before the Federal Communications Commission in the matter of implementing a broadband interoperable public safety network in the 700 MHz band.* Retrieved from PS Docket No. 06-229: 〈 http://fjallfoss.fcc.gov/ecfs/document/view?id=6520011227 〉

Peha, J. M. (2007a). Fundamental reform in public safety communications policy. *Federal Communications Law Journal , 59* (3), 517-546. 〈 http://www.law.indiana.edu/fclj/pubs/v59/no3/9-Peha.pdf 〉.

Peha, J. M. (2007b). How America's fragmented approach to public safety wastes money and spectrum. *Telecommunications Policy , 31* (10-11), 605-618.

Peha, J. M. (2005). *How America's fragmented approach to public safety wastes spectrum and funding.* Proceedings of the 33rd Telecommunications Policy Research Conference, 〈 http://web.si.umich.edu/tprc/papers/2005/438/Peha_Public_Safety_Communications_TPRC_2005.pdf 〉.

Peha, J. M. (2006). *The need for fundamental reform in public safety spectrum and communications policy.* Working Paper #15, October 2006, New America Foundation, Wireless Future Program, 〈 http://www.newamerica.net/files/WorkingPaper15_TVtoPublicSafety_Peha_FINAL.pdf 〉.

Peha, J. M., & Sutivong, A. (2001). Admission control algorithms for cellular systems. *ACM Wireless Networks , 7* (2), 117-125 〈 http://www.ece.cmu.edu/~peha/cellular_admission.pdf 〉.

Sesia, S., Toufik, I., & Baker, M. (Eds.). (2009). *LTE – the UMTS Long Term Evolution: From theory to practice* (1st ed.). Chichester, UK: Wiley.